

# Problems arising with genetic trend estimation in dairy cattle

B. BONAÏTI, D. BOICHARD, A. BARBAT, S. MATTALIA<sup>1</sup>

Station de Génétique Quantitative et Appliquée,  
Institut National de la Recherche Agronomique,

<sup>1</sup> Institut de l'Élevage,  
78352 Jouy en Josas Cedex

## Introduction

International rankings of dairy bulls are based on conversion formulae or on overall analysis of daughter yield deviations (DYD) of each sire \* country combination. Consequences of possibly biased estimations of within country genetic trend on these international rankings have never been studied. Such a bias would have probably little effect on within country selection, since choice of bulls and cows are mostly practiced among almost contemporary animals. Even if some bias exists, it may remain undetected within country, particularly if it is positive. However, this bias may strongly affect international comparisons: in a country which overestimates genetic trend, the young bulls will be favored in international comparisons. Therefore, each country should estimate its genetic trend as precisely as possible, and Interbull should give advice and define some rules in this aim. This work is a contribution to that purpose.

## I. Discrepancy between countries

Three observations indicate a large discrepancy between estimated genetic trends in France and in other countries.

### 1) Comparison with US pedigree

Since 1975, French AI studs have proven many bulls born from a US sire and a US dam evaluated in the USA. From 60 to 500 such bulls, born in USA or in France after embryo transfer, have been progeny tested in France each year. This sample of 1951 bulls provided an opportunity to compare expected genetic trend (estimated from pedigree proofs in USA) with realized trend in France (estimated from French proofs). ~~The most recent proofs available in June 1993 were used.~~

Table 1. Conversion factors from USA to France for milk, fat, and protein yield

Factor	Milk	Fat	Protein
a	-399	-28.9	-13.1
b	0.90	0.99	0.90

Each bull had a pedigree proof including either sire + dam information, or sire + maternal grand-sire (mgs) information. Before comparison with French results, these pedigree proofs were converted into French units with official formulae (table 1). The "b" factor was derived from within year standard deviations of proofs and DYD and assuming a genetic correlation equal to 0.9 between France and USA. The "a" factor was estimated by the difference between US proofs (x "b" factor) and French proofs of full sib bulls born since 1981 (Mattalia and Bonaïti, 1993). However, notice that the comparison of trends does not depend on the "a" factor.

We compared the realized trends in France and the expected ones from US pedigree proofs

(sire + dam or sire + mgs) for milk, fat, and protein yields. Proofs from the last 1992 run were considered in both countries. The pedigree trends were always significantly higher than the realized ones. Between 1978 and 1986 (birth years of bulls), the ratios of expected over realized trends were 2.65, 1.83, and 2.64 for milk, fat, and protein yields, respectively. The ratios were smaller with sire + mgs pedigree proofs which only partially account for the genetic trend of the dams: 1.92, 1.33, 1.94.

This discrepancy between expected and realized trends strongly suggests that genetic trend estimates are biased either in the USA, or in France, or in both countries. An overestimation of the b factors used in this study, or a decrease with time in the accuracy of dam proofs, are alternative hypotheses, which however do not seem sufficient to explain such a large discrepancy between countries.

## 2) Observations from EEC-Interbull group

EEC-Interbull group analysed DYD from four EEC countries (Germany, the Netherlands, Italy and France) and the USA using the method suggested by Schaeffer (1985). This study provided international proofs. A pedigree (sire + mgs) international proof might also be computed from these results for each bull. For bulls with US sire and mgs, the comparison between expected (pedigree proofs) and realized trends showed large discrepancies among countries. Moreover, as each proof combined all available information in this analysis, this discrepancy was likely to be still underestimated.

Such results may be explained by differences in the choice of dams: the weights of secondary traits (type, milk composition) or the accuracy in the choice of dams would have differently varied among countries. However, if this reason is insufficient to explain the observed differences, one must admit that genetic trend is biased in some national evaluations.

## 3) Variations of conversion factors

Some countries may observe that the "a" and "b" estimates vary with the bull samples. This is an indication of a possible problem in genetic trend estimation. If the "b" factor is assumed to be constant (theoretical "b") and if the genetic trend is overestimated in the exporting country or underestimated in the importing country, then the "a" estimate decreases when the average birth year of the sample increases. Similarly, with the same hypothesis on the genetic trend bias and if the "b" factor is estimated by Wilmink's method (1986), its estimate decreases when the number of birth years of bulls included in the analysis increases.

# II. Proposals to validate genetic trend estimations

Previous results showed that the estimation of genetic trend is likely to be biased in some countries, in spite of improvements given by the animal model and the BLUP methodology. Therefore, methods to detect biases on genetic trend are needed. Two methods are proposed and described below.

## 1) Comparison of results from first and all lactations data analysis

With an individual animal model evaluation system and when data of all parities are analysed, the estimation of genetic trend results from three different sources of information:

- \* the superiority of the animals selected as parents over their contemporaries (*i.e.* expected genetic trend)
- \* differences between performances of contemporary daughters issued from parents born in different years,
- \* and differences between contemporary performances of cows born in different years.

Obviously, the third source of information does not appear when only first lactation data are analyzed. Therefore, the estimation of genetic trend obtained from the analysis of first lactation data is

less likely to be biased, and, in any case, would not be biased by wrong age adjustment factors. As the genotype\*parity interaction is known to be small, both analyses are expected to provide similar estimates of genetic trend. When comparing the results of a first lactation data analysis with an analysis of all lactations, an important difference in estimated genetic trend would indicate the existence of a bias when using the second method.

### 2) Within-bull variation of DYD with production year

Daughter yield deviations (DYD) of bulls are average performances adjusted for dam breeding value and for all the effects included in the model, except daughter breeding value. Their expectation depends only on the bull, and they are, theoretically, independent of any environmental effect, and particularly on year of production. This property may be used to validate the estimation of genetic trend. Let us consider the following model to analyze the individual deviation of daughter  $k$ ,  $d_{ijk}$

$$d_{ijk} = s_i + y_j + e_{ijk}$$

with  $s_i$  being the effect of sire  $i$ ,  
 $y_j$  being the effect of the  $j$ th year of use of bull  $i$ ,  
 $d_{ijk}$  being the deviation of daughter  $k$  of sire  $i$ , obtained in the  $j$ th year,  
 and  $e_{ijk}$  being the error.

Usually,  $j=1$  or  $2$  (rarely  $3$ ) for the first crop of daughters, whereas  $j$  is greater than or equal to  $5$  (rarely  $4$ ) for the second crop of daughters.

When the estimate of genetic trend is unbiased, the year effect has a zero expectation and should not be significant. Alternatively, the year effect shows a decreasing or increasing trend when the estimate of genetic trend is underestimated or overestimated, respectively.

In practice, such an analysis does not require individual deviations, but only DYDs per sire and year of production, with appropriate weightings.

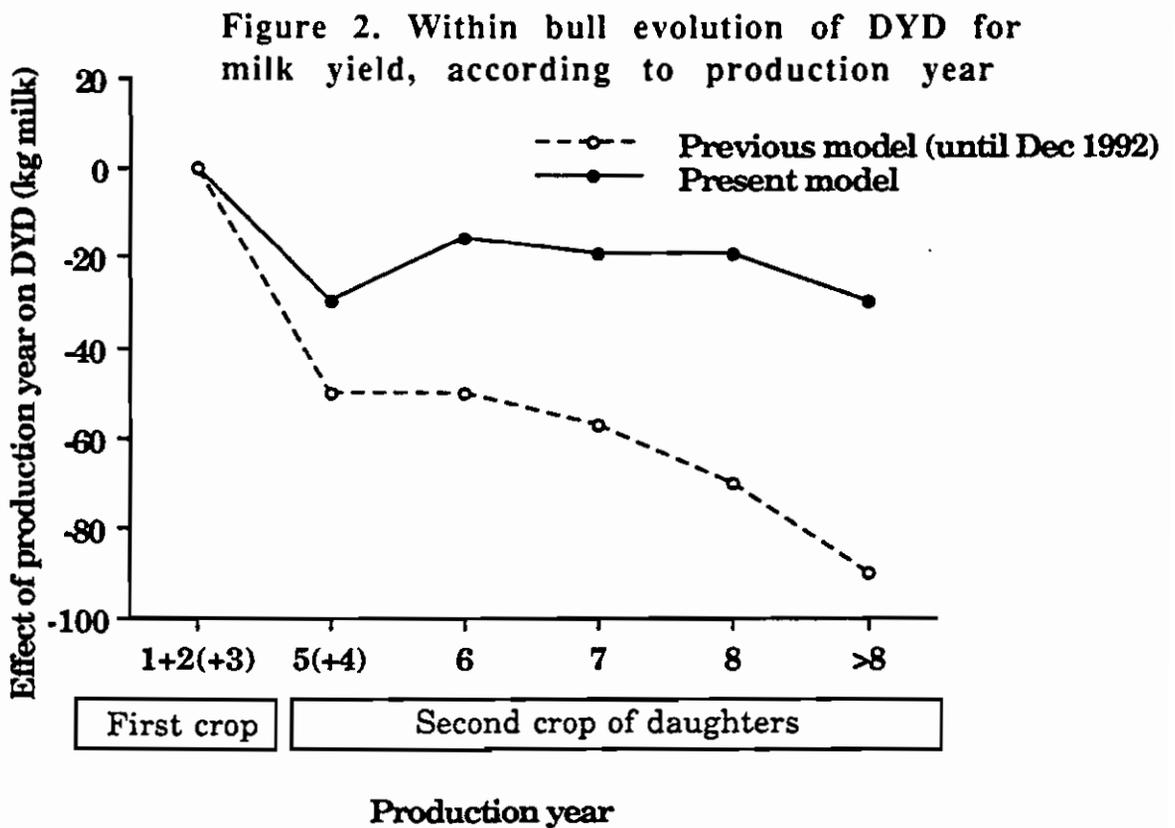
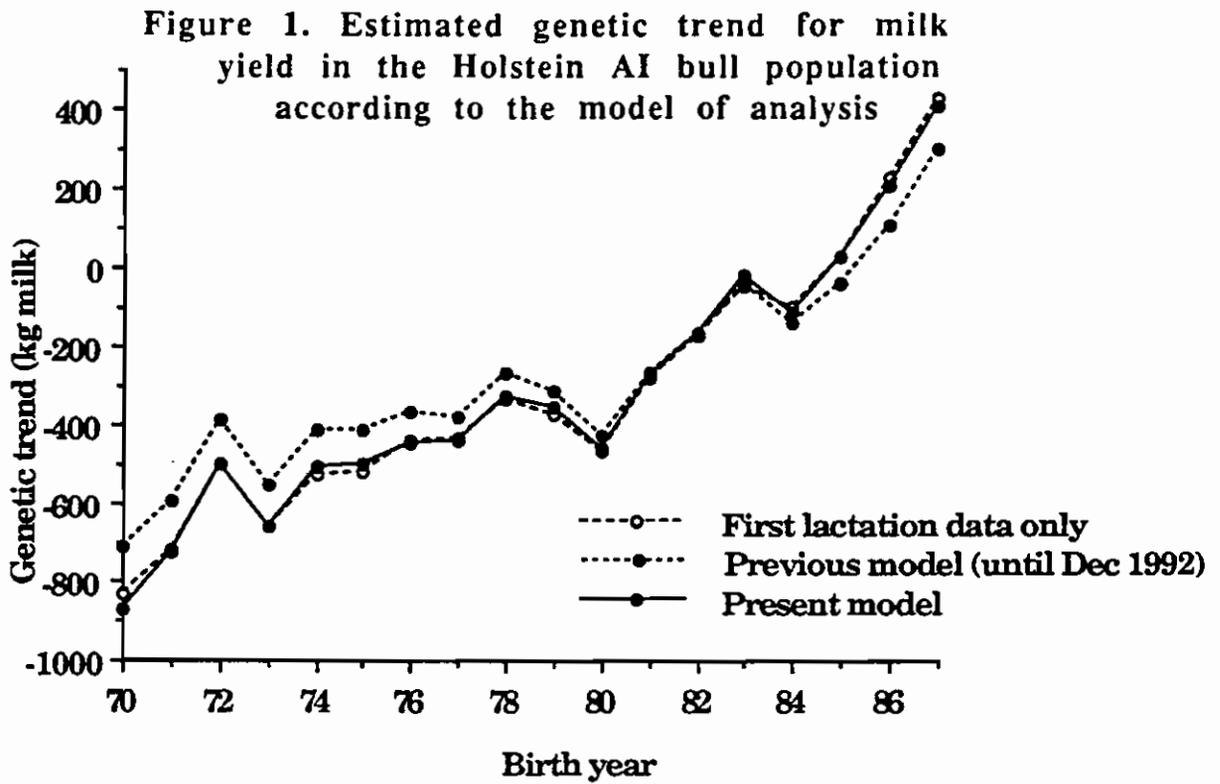
## III. Present status of the French Situation

### 1) Under estimation of the French genetic trend.

Since the animal model implementation in France in 1990, the genetic trend seemed to be underestimated in the Holstein population (Bonaiti & Boichard, 1990). Although they were not considered as proofs, some facts, already known three years ago, were quite alarming. In 10,000 herds raising animals of both Normande and Holstein breeds, environmental year effects increased more rapidly in Holstein than in Normande breed. Within sire and maternal grand-sire, bull proofs tended to decrease with birth year, in spite of a likely better value of the dam. To verify and quantify this intuitive feeling, both methods presented in part II were applied.

The results of an analysis of the first three lactations (official evaluation) were compared to those of an analysis restricted to the first lactation data. They do not completely agree since, in the population of Holstein AI bulls (>95 % Holstein genes), the estimated genetic trend was increased in the first lactation analysis by 22 kg milk per year (figure 1).

In the within-bull analysis of DYD, a negative trend was observed in the Holstein breed, as shown in figure 2. According to birth year of bulls, the DYD decreased by 40-60 kg milk between the first and second crops of daughters and by 0-10 kg per year within the service period. This trend was most pronounced for recent bulls (born after 1978), while it was of small magnitude for old bulls (born before 1975). Although these trends were quite small compared to the large variation in the environmental conditions and the level of the mates, they demonstrated clearly that the genetic trend was underestimated in the French Holstein population.



## 2) Evolution of the French dairy evaluation model

Dairy evaluation in France is based upon a repeatability animal model, including a genetic effect with groups for unknown parents, a cow permanent environmental effect, a herd-year effect, and some fixed effects defined within year and region. Until December 1992, these fixed effects included the effects of parity, calving month within parity, age at calving in first lactation, and calving interval for later parities.

Since March 1993, the model of analysis has been slightly modified. It included a preadjustment for heterosis and recombination losses, as described in Boichard *et al* (1992, 1993). However, the consequences of this preadjustment were very limited in practice, and the annual genetic trend was increased by 5 kg milk only in the Holstein AI bull population. The major change was the inclusion of an age effect within parity 2 and 3, in addition to the effect of age in first lactation. This age effect was of much smaller magnitude than in first lactation (500 kg and 300 kg milk between extreme classes in parity 2 and 3, against 1100 kg in first lactation) but was found to be not negligible. Finally, to avoid estimability problems, the effect of preceding calving interval in later parities was replaced by the effect of days dry. The levels of the age factor are given in table 2.

Table 2. Definition of the age factor levels (within parity, year and region) in the French evaluation system

Level	Parity 1	Parity 2	Parity 3
1	<= 23 *	<=36	<=48
2	24-26	37-39	49-51
3	27-29	40-42	52-54
4	30-32	43-45	55-57
5	33-35	47-49	58-60
6	36-38	50-52	61-63
7	>=39	>=53	>=64
8	Invalid or unknown	-	-

\* : age in months

## 3) Validation of the present estimation of genetic trend

The better adjustment for age generated a quite important increase in estimated genetic gain, rather consistent with that predicted with first lactation data only (figure 1). In the population of pure Holstein AI bulls born from 1978 to 1986, the estimated genetic trend increased from 54 kg milk per year to 73 kg, while the estimate from first lactation data reached 76 kg.

Furthermore, DYDs appeared to be more stable over time with the new model (figure 2), in spite of a small decrease (-15 to -30 kg milk) between both crops of daughters.

However, the discrepancy between realized genetic trend in France and expected genetic trend according to pedigree US proofs was only partially decreased. The ratios of expected over realized trends reached 1.99, 1.69, and 1.86 with the updated model (june 1993) instead of 2.65, 1.83, and 2.64 with the previous model (december 1992), for milk, fat, and protein yields, respectively. The figure 3 shows the realized and expected trends.

## IV. A potential source of bias: the preadjustment for age and parity

As already mentioned in part II, the genetic trend estimated from the analysis of all lactation data includes three sources of information : 1) the superiority of the animals selected as dam or sire over their contemporaries (*i.e.* expected genetic trend), 2) the difference between performances of contemporary daughters born from parents of various age, and 3) the difference between contemporary performances of cows of various age.

There is no known way to measure precisely the relative weights of these three sources of information. But some indications can be given. When the model includes a parity effect, the third source provides only little information because there is a close relationship between year of production, parity and birth year. In contrast, if data are pre-adjusted for parity or age at calving, as it is the case in practice in many evaluation systems, this third source becomes very important and possibly the main one, because there are a lot of possible comparisons of contemporary performances of cows born in different years. Consequently, the estimated genetic trend may be fully determined by the choice of the pre-adjustment factors for age at calving. This question led us to investigate the influence of pre-adjustment factors in the French situation.

In the French evaluation system, the age\*parity effects are estimated by region and year. On average in the recent years, they reached about 0, 1000 and 1300 kg milk in 1st, 2nd and 3rd lactations, respectively.

When data were preadjusted for these values (defined and computed precisely by year and region) and when the age\*parity effect was removed from the model, results were obviously unchanged for each effect remaining in the model, including breeding values. When data were preadjusted with over or under-estimated values and when the age\*parity effect remained in the model, results were still the same. But when the parity effect was removed from the model, the results were strongly modified. For example, when the differences between second and first lactations and between third and first lactations were increased by  $d=50\text{kg}$  (1050 kg milk instead of 1000 on average), and  $2d=100\text{kg}$  (1400 kg milk instead of 1300 on average), respectively, the estimated genetic trend was increased by 23 kg milk per year in Holstein breed. When considering different values of  $d$ , an almost linear relationship was observed between  $d$  and the estimated genetic trend: the annual genetic trend was biased by an amount close to  $0.4 d$ , even for very high values of  $d$ . Similar results were obtained in other breeds.

These results clearly show the large influence of age pre-adjustment on genetic trend estimation. As the bias may be very large, and as the strategy regarding age\*parity effect varies among countries (preadjustment, estimation in the evaluation, mixture of both systems when the contemporary groups distinguish heifers from older cows...), this bias may be an explanation of the discrepancy observed among countries.

## CONCLUSION

In this paper, we have pointed out a large discrepancy between genetic trends estimated in France and the USA. Some facts (observations from EEC-Interbull groups, unexplained variations of conversion factors) suggest that the problem may not be limited to these two countries. To overcome this problem, two simple methods have been proposed to validate the estimation of genetic trend within country.

These methods, applied to the French data, demonstrated that the genetic trend was downwards biased in France until December 1992, and were used to validate some changes in the present analysis model of the evaluation system. However, in spite of the quite large increase in estimated genetic gain in the French population resulting from these changes, the discrepancy between France and the USA was only partially decreased.

Such an analysis could be realized by other countries, particularly those exporting or importing genes. Each country could 1) compare the results of its official evaluation with the results of an analysis restricted to the first lactation data, and 2) analyse the within-bull evolution of DYD over time. Moreover, because of the extreme sensitivity of the estimated genetic trend to age adjustment factors, we would like to recommend to estimate these age (or parity) factors in the evaluation model. At least, the estimation conditions of the pre-adjustment factors should be described precisely.

From our view point, these additional investigations could make national evaluations more reliable and more comparable to each other, and consequently could greatly improve the accuracy of international comparisons, which is the natural objective of the INTERBULL committee.

## REFERENCES

- D. Boichard, B. Bonaïti, A. Barbat, 1992. Effect of the Holstein crossbreeding on genetic evaluation of the French Black and White cattle population. Interbull meeting, Neufstift, Austria, June 7-8, 1992. *Interbull bulletin* No 7.
- D. Boichard, B. Bonaïti, A. Barbat, 1993. Effect of Holstein crossbreeding in the French Black and White cattle population. *J. Dairy Sci.*, **76**, 1157-1162.
- B. Bonaïti, D. Boichard, 1990. Benefits from animal model evaluation of dairy cattle in France. *Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, Scotland, XIII*, 364-373.
- S. Mattalia, B. Bonaïti, 1993. Use of full sib families to estimate the "a" coefficient of conversion formula between countries. Interbull, Aarhus, Denmark, August 19-20 1993.
- R.L. Schaeffer, 1985. Model for international evaluation of dairy sires. *Livest. Prod. Sci.*, **12**, 105-115.
- J.B.M. Wilmink, A. Meijering, B. Engel, 1986. Conversion of breeding values for milk from foreign populations. *Livest. Prod. Sci.*, **14**, 223-229.

Figure 3. Comparison of French proofs with US pedigree proofs for 1951 Holstein bulls  
Milk yield

